Contextual Cues in Sensory Testing

Jamie Zumach[1*], Dr. Christopher Simons (Laboratory Supervisor)
[1] Laboratory member. Food Science and Technology Dept. The Ohio State University, Columbus, OH 43210
*Jamie Zumach (zumach.1@osu.edu)

Abstract

Since the early 1900s, sensory analysis has provided food corporations with key information about the demands of consumers. However, a lack of development in the field of sensory science and unreliability of current methods has contributed to high failure rates of new food products. Traditional testing booths lack ecological validity, which likely contributes to data with poor quality. One of the few attempts at remedying the issue has been to allow the consumer to interact with the product in a more realistic environment than the laboratory setting with at-home testing. Despite the valuable results and insights this method may provide, it can also increase error due to the potential lack of standardization and control. The purpose of this study was to investigate whether providing relevant contextual information through an immersive technologies environment (e.g. virtual coffee house) alters the consumer's hedonic perceptions of black coffee compared to traditional in-booth evaluations. Panelists were given five black coffee samples in both testing settings, and rated their liking for each product using a nine-point hedonic scale. Preliminary data revealed that the virtual reality (VR) environment was associated with a significant difference in liking between black coffees, whereas the standard booth was not. As a result, the preliminary results suggest that the virtual reality environment provides more sensitive data by decreasing variability in liking scores, and that context impacts how different coffees are perceived. In order to collect more data to support these results, sensory testing will be run with approximately 50 individuals in both sensory testing settings. The final outcome of this set of tests will assist in determining further sensory

research involving the VR lab and its efficacy as an advantageous sensory testing environment.

Key words: sensory analysis, immersive technologies, standard booth, hedonic scale, liking scores

Introduction

The roots of sensory analysis began with trade of goods, utilized in order to determine the quality of a product and furthermore its appropriate price. As a result, these practices were the beginning of standardized product grading, which has since transformed and modernized to become what is known as sensory analysis. The foundation of modern sensory analysis has its roots in the 1940s to mid 1950s, when the U.S. Army Quartermaster Food and Container Institute began utilizing sensory analysis to determine the likability of food products designed for the armed forces. Furthermore, in the 1949 the U.S. Army Quartermaster Laboratory developed the Hedonic scale. Due to the results the government received from their sensory evaluation, the food industry began implementing sensory analysis as well into their operations. Companies that incorporated sensory evaluation into their research and development (R&D) processes profited from the data received through testing (Clark and others 2009). Due to this success, eventually all large food corporations integrated sensory analysis as an essential discipline in their R&D programs.

Although sensory analysis has proven shown to provide valuable insight into the desires and needs of consumers, the current methods utilized have revealed some crucial limitations. Despite the long history surrounding hedonic assessment and often the completion of comprehensive sensory testing, many new products still fail when they reach market. Estimations place new product failure at upwards of 70-80% in modern times (Stanton 2013). A potential explanation is that the flawed sensory results are the repercussion of a lack of relevant context presented in the traditional sensory environment, leading to unreliable and

highly variable data. Consequently, the food industry has considered other alternatives to the traditional methods of sensory analysis. One of such variations is the central location test (CLT), which involves product testing that can occur in various settings. The flexibility in location may allow for greater consumer recruitment, as well as the potential for incorporating relevant contextual information into the testing environment. The testing process can be controlled, and the sensory staff has oversight over the operation. However, as with traditional sensory booths, the participants analyze a small portion of the product and thus have limited exposure. Another alternative is home use tests (HUT), in which the consumer conducts the test in their home over an extended period of time. The consumer is able to encounter the product in situations reflective of their daily lives, and develop an opinion following multiple uses. Due to the incorporation of pertinent context, the data received may be a more valid or accurate portrayal of consumer opinion. However, HUTs are expensive to conduct and also relinquish control over sensory conducting procedures to the consumer (Lawless and Heymann 2010). As a result, companies may be reluctant to administer HUTs despite the valuable data they can provide. Consequently, although alternatives to the traditional methods of sensory analysis have been introduced, many corporations still rely on the standard techniques.

Hypothesis

In traditional sensory testing methods, participants are isolated in booths, which present little context, and consume portioned amounts of a product. Although these methods of sensory analysis have proven useful in the ability to manipulate and control the setup of testing, the data gathered can often be unsound and incomplete. As a result, we believe that context is influential in the formation of the consumer perception and opinion, and thus are interested in its role in and impact on sensory analysis. Furthermore, we hypothesize that the liking data

obtained from the VR environment will be more sensitive and reliable in comparison to the traditional sensory booth due to the influence of relevant context.

Purpose and Research Objectives

The purpose of the study is to draw comparisons between the data collected from a standard sensory booth and the VR environment created through use of immersive technologies. The research objectives are the:

1. Comparison of data obtained from a liking questionnaire given in both a standard sensory booth and the VR environment on the topics of preference, resolution and scale.

2. Determination of the impact of context in the evaluation of liking of coffee samples, and the discernment of which environmental cues are most relevant during testing.

3. Establishment of plans for further research and potential development of sensory testing methods to be utilized in industry

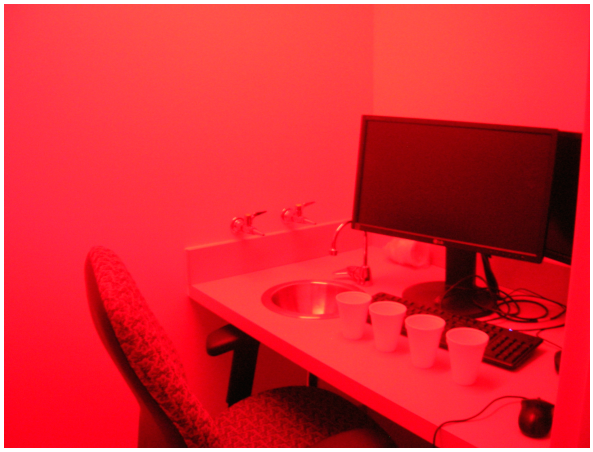
Methods and Materials

Thirteen participants who frequented coffee shops twice a week and drink black coffee were recruited for the study. Each participant gave written consent to participate. Coffee samples were prepared from four brands: Starbucks Pike Place Blend (Seattle, WA), Maxwell House (Kraft, Glenview, IL), Stauf's Ethiopia Moka Harras (Columbus, OH), and Douwe Egberts (DE) Java Coast Espresso (Amsterdam, The Netherlands). Each of these coffees were chosen for their similarity in roast and difference in price, as well as their availability for purchase as whole bean. The coffees were stored in the freezer prior to use, and were ground with a Capresso 560 Infinity Burr Grinder on a "medium" grind setting before brewing. For the preparation of the coffee samples, 50 grams of grounds was weighed and then placed into a Sunbeam coffee maker (Boca Raton, FL), which was lined with a paper filter. An additional fifth coffee sample, utilizing the DE brand, was prepared with 1.5 times the amount of grounds

(75 grams), to produce a stronger brew. Then 2,100 grams of water was poured into the coffee maker water reservoir and the coffee maker was turned on. After the five coffee batches were brewed, each were poured into Choice 2.5 liter air pots (Lancaster, PA) from which the samples were taken. The coffee samples were dispensed into either Styrofoam cups or ceramic mugs, depending on the testing environment. All samples were labeled with a randomized three digit number.

Prior to the testing, the sensory evaluation rooms (shown in Figure 1) were prepared with the appropriate materials and sensory stimuli. A traditional sensory testing booth in the basement of the Parker Food Science Building on the Ohio State University campus was prepped and coffee samples were served under red lighting and in Styrofoam cups. In the adjacent room, which was the Immersive Technologies Laboratory, auditory, visual, and olfactory elements were prepared. Video footage and audio recorded at the One Line Coffeeshop in Columbus, OH was displayed on nine LCD screens and played through speakers located in the room. Additionally, an olfactometer released cinnamon roll aroma (Givaudan Flavors, Cincinnati, OH) into the testing environment. As well, ceramic mugs were utilized for the presentation of the coffee samples to the participant.

A



B

Figure 1. (A) Traditional sensory booth set up (psychophysiology lab room) utilized in testing. (B) Virtual reality environment set-up utilized in testing.

During the sensory evaluation, a participant entered one room and was presented with the coffee samples and water on a tray, and a ballot. The order of the samples on the tray was counterbalanced in order to prevent potential order effects. The participant would taste each coffee sample, rinsing with water in between, and rate their liking of each on a nine-point hedonic scale (similar to that shown in Fig. 2). After they were finished, the process would be repeated in the other testing setting. At the completion of the test, the participant was given a $10 gift card.

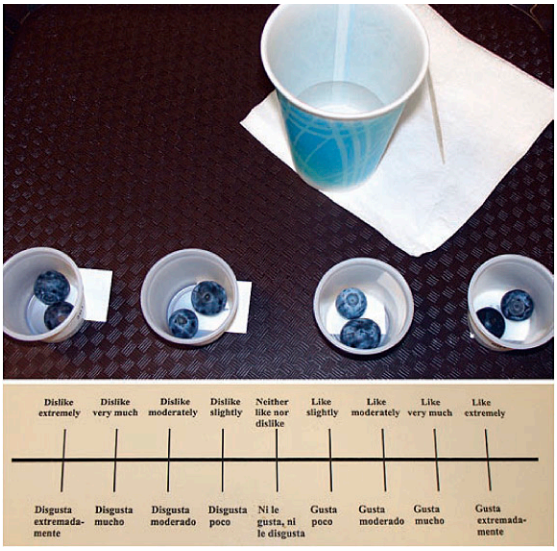


Figure 2. Example of nine-point hedonic scale[1]

Results and Discussion

Results from ANOVA indicated a difference in the data obtained from the traditional sensory booth and the VR environment on three points: preference, resolution (presence or lack of significant differences in liking scores between samples), and scale (usage of the values on the nine-point hedonic scale). The preference order is important because variations between testing conditions may indicate that context influences and assists with altering the liking scores of a

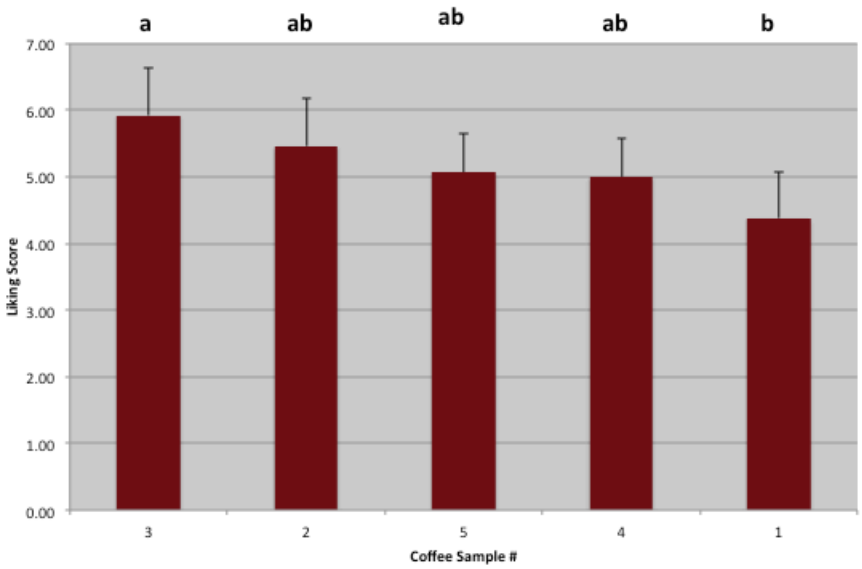
product. Additionally, a comparison of resolution (or lack there of) between samples in the two environment can identify that context promotes greater sensitivity of participants to differences between products (and thus greater variation in liking scores). Lastly, analysis of the usage of the hedonic scale for both environments provides information about how context affects the ability of participants to determine product differences and furthermore, distinguish preferences.

Preference:

Figure 3. Preference determined from liking scores in virtual reality and traditional testing environments.

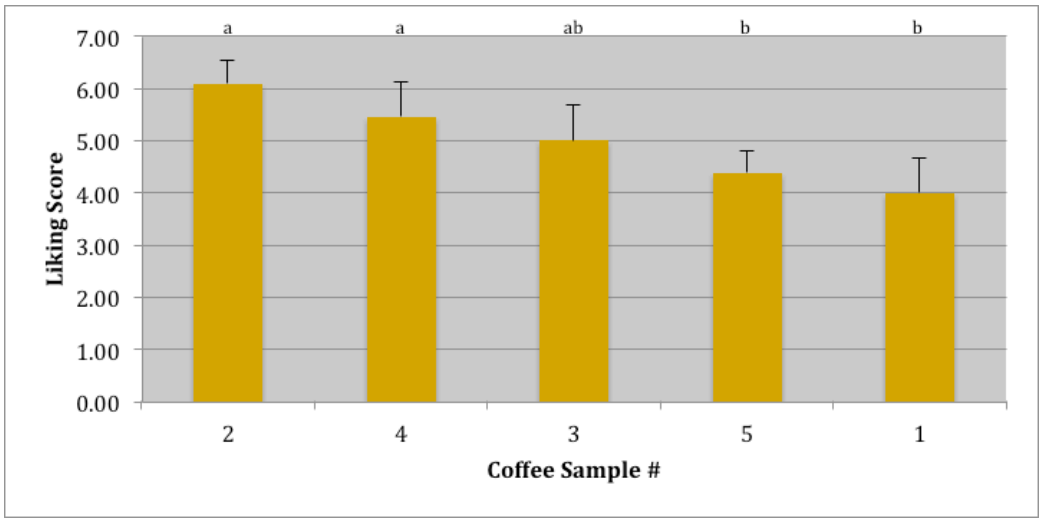
In figure 3, a differing preference order of coffees between each testing condition can be identified. For the VR environment, Starbucks was the most preferred sample with a liking score of 6.08, followed by Stauf's (5.46), DE 1.5 (5.00), DE (4.38), and lastly Maxwell House (4.00). In comparison, in the traditional booth, DE 1.5 was the most liked sample with a score of 5.92, followed by Starbucks (5.46), DE (5.07), Stauf's (5.00), and lastly Maxwell House (4.38). These results support the idea that premium coffee (i.e. Starbucks and Staufs) may be more preferred in an environment that provides context similar to that which is encountered during consumption.

Resolution:



Coffee Sample #:
1: Maxwell House
2: Starbucks
3: Douwe Egberts 1.5
4: Stauf's
5: Douwe Egberts

Figure 4. Significant difference of coffee samples based on liking in traditional testing environment.



Coffee Sample #:
1: Maxwell House
2: Starbucks
3: Douwe Egberts 1.5
4: Stauf's
5: Douwe Egberts

Figure 5. Significant difference of coffee samples based on liking in virtual reality environment.

In Figures 4 and 5, the statistical differences (determined from ANOVA; $p<0.05$) between the samples are displayed for the traditional testing environment and the VR environment. The graph is arranged in decreasing liking score from left to right, and standard deviation bars are associated with each score. For the traditional environment, a significant

difference in liking scores can only be determined for samples three and one. For the virtual reality environment, a significant difference in liking scores is noted between samples two and five, two and one, four and five, and four and one. As a result, more significant differences have been established between coffees in the VR scenario, and thus greater statistical power was associated with data obtained from the VR environment.

Scale:

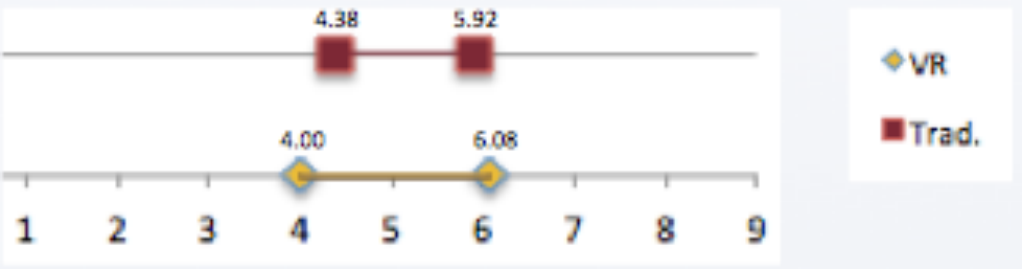


Figure 6. Usage of hedonic scale of liking for coffee samples in virtual reality and traditional testing environments.

Figure 6 compares the use of hedonic scale by participants in the VR and traditional environment during the study. For the traditional environment, the hedonic scale utilized was between 4.38 and 5.92. In comparison, the scores from the VR environment illustrate greater use of the scale, from 4.00 to 6.08. As a result, the contextual cues provided by the VR environment may assist with the perception of liking differences between coffees leading to a wider use of the hedonic scale. The utilization of a wider range of scale allows for identification of more significant differences between samples, and a more clear picture of the consumer preference among the samples.

Conclusion

In the study, the objective was to examine the hypothesis that the provision of contextual cues through the use of immersive technologies would lead to more sensitive and reliable data than the traditional method in sensory analysis. Through examination of the results obtained and

data analysis through ANOVA, differences were determined between the traditional environment and the VR environment for preference, resolution, and scale. The testing environment was found to impact the order of preference, and more significant differences established in the VR environment signified a greater sensitivity to coffee samples in that setting. Additionally, a wider range in liking for the VR situation was determined, indicating that engagement may be greater in the VR environment.

Although the hypothesis the study was examining was supported by the acquired results, the trends discussed above are based off sample size of thirteen. As a result, these conclusions may not be accurate or stable amongst a larger pool of participants. In order to strengthen the accuracy of the data, a full study with fifty participants would need to be completed and the results analyzed with ANOVA. Additionally, a follow-up study would be completed with those fifty participants a month following the initial test in order to test the reliability of the data. Based on the results from the comprehensive test with fifty participants, more accurate and reliable conclusions may be made about the influence of context in sensory analysis. Furthermore, these results could then be utilized by the food industry to develop more reliable and sensitive methods of sensory evaluation in order to more accurately gauge consumer liking and thus decrease the occurrence of product failure.

Acknowledgements

I would like to give special thanks to the project advisor, Dr. Simons, and the members of the lab team who assisted with the project (Kacey Smith, Elizabeth Green, Ronald Bangcuyo, Elka Del Portal, Alex Pierce, Samuel Subler, and Gretchen Guttmann). An additional thanks to One Line Coffee for assistance with the project through allowing video footage and audio recording to be taken in their shop.

References

1. Bremer V., Crisosto G., Molinar R.H., Jimenez M.J., Dollahite S., Crisosto C.H. 2008. San Joaquin Valley blueberries evaluated for quality attributes. California Agriculture [serial online]; 62(3): 91-96. Available from: http://ucanr.org via the Internet. Accessed 2014 Feb 19.
2. Clark S., Costello M., Drake M., Bodyfelt, F. 2009. Chapter 1: History of Sensory Analysis. In *The Sensory Evaluation of Dairy Products*. Springer.p. 1-6.
3. Lawless H.T., Heymann H. 2010. Chapter 15: Consumer Field Tests and Questionnaire Design. In *Sensory Evaluation of Food: Principles and Practices*. Springer. P. 349-378.
4. Stanton J. 2013. Market View: The Return of Truly New New Products. Food Processing. Available from: http://www.foodprocessing.com/articles/2013/increased-new-products/. Accessed 2014 Mar 25.